AD_____

Award Number: W81XWH-04-1-0277

TITLE: Quantitative in Situ Assessment of the Somatostatin Receptor in Breast Cancer to Assess Response to Targeted Therapy with 111-in-Pentetreotide

PRINCIPAL INVESTIGATOR: Gina G. Chung, M.D.
                                                David Rimm, M.D., Ph.D.

CONTRACTING ORGANIZATION: Yale University
                                                    New Haven, CT 06520

REPORT DATE: May 2008

TYPE OF REPORT: Final Addendum

PREPARED FOR: U.S. Army Medical Research and Materiel Command
                            Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
                                            Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01-05-2008	Final Addendum	1 May 2007 – 31 Mar 2008

4. TITLE AND SUBTITLE

Quantitative in Situ Assessment of the Somatostatin Receptor in Breast Cancer to Assess Response to Targeted Therapy with 111-in-Pentetreotide

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-04-1-0277

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Gina G. Chung, M.D.

David Rimm, M.D., Ph.D.

E-Mail: gina.chung@yale.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Yale University
New Haven, CT 06520

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Somatostatin (SST) is a peptide hormone implicated in the growth and progression of cancers and SSTR2 is the predominant receptor subtype expressed in breast cancer. We hope to study the pattern of expression and clinical significance of SSTR2 levels in breast cancer. We have developed an algorithm called AQUA that can assess protein expression on tissue microarrays (TMA) based on molecular co-localization techniques. Our results show that SSTR2 is localized predominantly to the malignant cells although also in vessel/lymphatic elements. Although expression was not significantly correlated with survival on our TMA, it did appear to be overexpressed compared with benign breast tissue. A vessel compartment has been developed using a multiplexing protocol for co-localization of SSTR2 to tumor and endothelium concurrently. Cell line controls have also been developed to generate a reference standard curve and to use as a normalization feature. Whole sections analysis with SSTR2 also show that while heterogeneity of expression is present, it is modest.

15. SUBJECT TERMS
breast cancer, quantitative analysis, tissue microarray, somatostatin

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** (include area code)
U	U	U	UU	41	

Table of Contents

Introduction

Somatostatin (SST) is a peptide hormone that inhibits the release of various hormones and growth factors. The receptors are also expressed in numerous tumors, with SSTR2, the predominant subtype expressed in breast cancer. Although there are some data for inhibitory effects of SST analogues in breast cancer, to date, small clinical trials of these agents have not been successful, perhaps in part because SST_R status prior to treatment was minimally investigated and varied in these studies. Until recently, SST_R expression has been performed by labor intensive methods such as autoradiography and RT-PCR *in vitro* and scintigraphy *in vivo*. We have developed a series of algorithms called AQUA that can assess protein expression on tissue microarrays (TMA) based on molecular co-localization techniques. Our automated analysis involves immunohistochemistry (IHC) combined with semi-automated acquisition and analysis of compartmentalized, quantitative, continuous scores which removes the inherent subjectivity of standard pathologist-based scoring systems. Our proposal further characterizes the expression and clinical significance of SST_R2 using large cohort breast cancer TMAs and outlines a means to translate and normalize the AQUA measurements from TMAs to whole section for clinical applications. In this manner, we hope to direct the development of targeted therapies to SSTRs more rationally.

Body

Task 1. Characterize SSTR2 expression in a breast cancer TMA

The goals of this aim have been completed. In summary, using AQUA analysis of SSTR2 expression on multiple fold redundant large cohort breast cancer TMA, SSTR2 stained predominantly in the invasive tumors in a membranous pattern, and to a lesser extent, in the stroma and vascular/lymphatic structures as well. And although in our cohort of patients, SSTR2 expression did not correlate significantly with survival, the clear overexpression of SSTR2 in tumors (as compared with benign breast epithelium)and the predominant tumoral rather than stromal localization suggest that future studies of SSTR2 as a homing target for labeled somatostatin analogues may be an effective strategy.

Task 2. Translating TMA-based AQUA algorithms to whole sections

Whole section analysis of ER from multiple slides/blocks of breast cancer has been completed and has recently been published (1). This showed good concordancy among blocks from the same specimen using AQUA or pathologist-based binary measurements but poorer concordancy when continuous AQUA scores were used.

The same breast cancer whole sections have been stained with SSTR2 (same antibody used for TMA experiments) and AQUA analysis of the entire sections have been performed using similar methods to what was described in the above manuscript. This analysis showed that whereas there is heterogeneity of

scores within a section, the variability was not as significant. Scatter plots of AQUA scores showed a relatively low level of variance for both low SSTR2 and high SSTR2 expressing cases (Figure 1). These results suggest that a relatively few representative areas of tumors could be sampled (either as TMA redundancy or as separate areas within a section) to represent SSTR2 expression within a tumor.
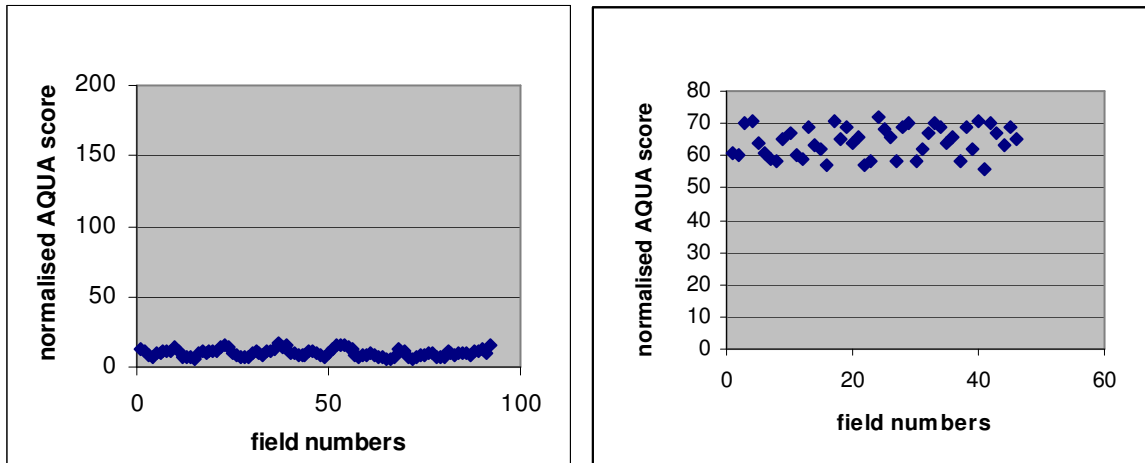

Figure 1. Scatter plots for SSTR2 scores on whole sections of breast cancer

Task 3. Conversion of AQUA to a protein concentration

In order to construct a standard curve for SSTR2 to serve as a reference point for conversion of AQUA scores of SSTR2 in tissue to protein concentration, we attempted to construct a serial dilution of SSTR2 peptide arrays in a matrix such as collagen. This however was quite problematic secondary to a variety of reasons, predominantly related to difficulties in constructing a peptide plug that was still stainable and readable by the AQUA protocol. ELISA measurements of these cell lines have also been problematic as concordance with AQUA scores of cell line arrays (see below) has been difficult to reproduce. We suspect this may be related to variances in tissue culture conditions when performing both procedures.

We therefore have begun to use established breast cancer cell lines as a reference measure. 'Boutique' arrays are small arrays of 10-20 cancer cell lines processed into a cell line microarray (a procedure our lab has utilized in previous experiments involving harvesting the cell lines, formalin fixation, resuspension, then conventional paraffin embedding) have been constructed. An example of the range of SSTR2 AQUA scores from these boutique arrays has been previously reported. These arrays can then be added on to tissue arrays or whole sections adjacently on the same slide or at least stained side-by-side with the experimental slide. In this manner, slide-to-slide normalization is possible for comparative studies. In the future, we propose that all SSTR2 measurements with AQUA in tissue be performed with these cell line controls. All whole section analysis studies as well as final TMA analysis have been performed with this normalization procedure.

Key Research Accomplishments

1. SSTR2 is expressed predominantly in the membrane compartment of breast tumor cells based on in situ AQUA measurements but is also present in stromal elements including apparent vessel compartments. Although SSTR2 is overexpressed in malignant breasts tumor cells compared with normal breast epithelial cells and was associated with several standard breast cancer prognostic parameters, it was not associated with survival.
2. An endothelial compartment can be constructed with AQUA using specific antibodies (e.g. CD31) that is prognostic of outcome and associated with breast cancer parameters (e.g. tumor size). This compartment can be used as another co-localization parameter such that quantitative measurements of SSTR2 within tumor and vessel compartments can be simultaneously obtained.
3. SSTR2 can be stained and analyzed by AQUA on whole tissue sections following the algorithms established by estrogen receptor. Grids of the whole sections can be placed 'virtually' on the section such that the entire section can be analyzed rapidly and efficiently. ER heterogeneity was most marked with continuous measurements. Further analysis of SSTR2 stained whole sections show that heterogeneity of expression was fairly minimal and that analysis of only a few (2-3) fields were representative of whole sections.
4. Boutique cell line arrays have become more readily accomplished. SSTR2 measurements in consistently designed boutique arrays are more consistent although correlations with ELISA measurements were inconsistent.
5. Clinical protocol studying SSTR2 expression by three different methodologies in patients with sarcomas or breast cancers

Reportable Outcomes

1. Somatostatin Receptor Subtype 2 (SSTR2) Expression in Breast Cancer. Ghosh S, Zerkowski M, Camp RL, Rimm DL, Murren J, Chung GG. Submitted.

2. Quantitative Analysis of Estrogen Receptor Heterogeneity in Breast Cancer. Chung GG, Zerkowski MP, Ghosh S, Camp RL, Rimm DL. Lab Invest, 87(7):662–669, 2007.

3. Decreased Expression of Somatostatin Receptor Subtype 2 (SSTR2) in Pancreatic Adenocarcinoma is Associated with a Poor Prognosis. Ghosh S, Yoon H, Harigopal M, Salem RR, Psyrri A, Camp RL, Rimm DL, Burtness BA, Murren JR, Chung GG. Cancer Invest, in press.

4. Biomarker validation: *in situ* analysis of protein expression using semi-quantitative immunohistochemistry-based techniques. Sullivan CO, Chung GG. Clinical Colorectal Cancer, 7(3):172–177, 2008.

5. Yale University HIC protocol 12513: Characterization of Somatostatin Receptor Expression in Sarcomas and Breast Cancer

Personnel receiving pay from research effort
   1. Maciej Zerkowski
   2. Sriparna Ghosh
   3. Catherine Sullivan

Conclusions

We have begun a systematic analysis of the expression of the SSTR2 in breast cancer using our automated analysis methodology which allows rapid, reproducible, quantitative measurements of in situ protein expression on tissue arrays. Our results show that SSTR2 is expressed in a graded fashion in a large proportion of breast cancers, is expressed predominantly within tumors and less so in stromal elements, and that it is mostly expressed in the membrane compartment of tumors. Although expression was not significantly correlated with survival on our TMA, it did appear to be significantly overexpressed in malignant breast epithelium compared with benign breast tissue. These results have now been reproduced in multiple fold, large cohort TMAs with several different antibodies. Because SSTR2 have been implicated in tumor angiogenesis and because our initial results suggested SSTR2 localized to tumor associated microvessels, we have also initiated work on creating a vessel compartment with AQUA. This has been readily accomplished using similar AQUA algorithms and a CD31 antibody and shows that an AQUA-based microvessel area is feasible, that it is associated with survival and other prognostic parameters in breast cancer, and that co-staining with other markers (e.g. VEGF, SSTR2) using a multiplexing protocol is feasible. Furthermore, cell line controls have been developed into "boutique array" with known relative levels of SSTR2 to serve as inter-slide normalization measures. Using ER as a prototype biomarker in breast cancer, we have translated the AQUA methodology to whole sections and adapted this protocol to analyze SSTR2 on whole sections.

References

1. Quantitative Analysis of Estrogen Receptor Heterogeneity in Breast Cancer. Chung GG, Zerkowski MP, Ghosh S, Camp RL, Rimm DL. Lab Invest, 87(7):662-669, 2007.

Appendix

1. Biomarker validation: *in situ* analysis of protein expression using semi-quantitative immunohistochemistry-based techniques. Sullivan CO, Chung GG. Clinical Colorectal Cancer, 7(3):172-177, 2008. (See attached PDF-23 pages)

2. Quantitative Analysis of Estrogen Receptor Heterogeneity in Breast Cancer. Chung GG, Zerkowski MP, Ghosh S, Camp RL, Rimm DL. Lab Invest, 87(7):662-669, 2007. (See attached PDF-8 pages)

Biomarker validation: *in situ* analysis of protein expression using semiquantitative

immunohistochemistry-based techniques

Catherine A.W. Sullivan,[1] Gina G. Chung[1*]

[1] Yale University School of Medicine, Yale Cancer Center, Section of Medical Oncology, 333 Cedar St, New Haven, CT, 06520

* Corresponding author
Yale Cancer Center
Medical Oncology
333 Cedar Street
P.O. Box 208032
New Haven, CT 06520

Email: gina.chung@yale.edu
Phone: 203-737-1569
Fax: 203-785-3788

Abbreviated title: quantitative automated analysis

Key words: biomarkers, protein expression, microarray, immunohistochemistry, automated

analysis, molecular profiling

ABSTRACT

Biomarker driven cancer research has become common in current literature. Much of this is driven by the increase in genomic and proteomic high throughput technologies that have increased our knowledge but has also produced a plethora of data with unclear clinical significance. Immunohistochemistry based assessment of protein expression is a natural validation method of expression profiling data that is easily performed on tissue samples collected prospectively or from archived samples. Coupled with tissue microarray technology and the increasing numbers of available automated, quantitative systems to read these arrays, we now have an efficient manner to validate biomarkers for prognostic and predictive capabilities as well as for identification of drug development targets.

INTRODUCTION

Although traditional clinicopathologic measurements of cancers such as stage and tumor grade are extremely helpful in determining the prognosis of a patient and to guide us in treatment recommendations, they are imperfect and likely lead to frequent under- and over-treatment. Thus, there has been great interest in the identification and validation of new biomarkers to more precisely guide us in these clinical decision-making processes and also to help us better understand tumor biology and aide in drug development strategies.

With the advent of high-throughput genomic technologies such as cDNA microarrays that allow the simultaneous analysis of thousands of genes from biologic samples, we now have a multitude of studies that have identified potentially useful candidate biomarkers for the above purposes.[1-4] Whereas these approaches have affirmed the complexity of cancer, the translation of these promising gene profiling studies to clinically useful tests have been difficult. Part of this difficulty is self-imposed as expression arrays can sometimes produce erroneous data with false negative and positive results.[5] In addition, these studies for the most part rely on obtaining fresh, frozen, non-fixed tissue making widespread clinical diagnostic applications more challenging.

In addition to refining these technologies, it is now increasingly recognized that confirmation of these expression array data is critical. A variety of methods is available and has been utilized. At the RNA level, methods such as Northern blots, real-time RT-PCR, and *in situ* hybridization (ISH) have been used effectively but can have technical challenges and is not always easily implemented.[6] Furthermore, a more likely relevant validation may be to study the expression

level of the biomarkers at a protein level such as with immunohistochemistry (IHC). Studies have shown that there is often discordance between levels of nucleic acids and proteins implying that the study of both measures is important. In any case, validating the functionality and clinical relevance as well as ensuring reproducibility and broad applicability of these biomarkers is an important priority.

Methods of Gene Expression Validation in Tissue

Several methods of confirming gene expression have been described. In general, RNA-based evaluations such as RT-PCR and Northern blot experiments are a vital initial step in the process but face a multitude of limitations. These include optimizing the stability of RNA, requirements for large amounts of material from often fresh/frozen tissue (esp. with Northern blots), separating tumor from non-tumor tissue (with procedures like laser capture microdissection, LCM), cost, and time requirements.[6, 7] Furthermore, these studies do not maintain the integrity of the tissue. Alternatively, *in situ* studies allow us to look at genes and proteins within the context of tissue morphology. With *in vitro* models, we can infer that the results are what may be occurring in the cells of interest. However with *in situ* studies we can actually investigate the levels of nucleic acids/proteins on an individual cell (and sometimes sub-cellular) level, while anatomically preserving the tumor microenvironment.

With the development of *in situ* hybridization (involving the use of a labeled complimentary DNA or RNA strands to localize a specific DNA or RNA sequence of interest), identification and localization of chromosomes and chromosomal integrity can be evaluated. In addition, with

careful normalization standards, quantification of amplified gene signals and mRNA transcripts of interest can be analyzed as evidenced by fluorescence *in situ* hybridization (FISH) studies of Her2/neu in breast cancer.[8, 9]

At the protein level, the majority of validation assays have used immunoblots, ELISA assays, tissue lysate arrays (lysates from cells collected by LCM arrayed onto nylon coated slides), and immunohistochemistry.[10-12] Although methods such as ELISA may be more truly quantitative as measurements are reported in absolute concentrations, IHC is however the only method that maintains morphologic information. The tissue microarray (TMA) methodology represents an excellent means of using both ISH and IHC techniques in a large set of tissue samples.[13] Some of the most commonly used *in situ* methods are outlined in Table 1.

Immunohistochemical Analysis

IHC has been used as a standard diagnostics tool in pathology since the late 1970's. It involves a series of semi-standardized steps. These include the removal and processing of a tissue sample from a donor, antigen retrieval, application of a primary antibody specific to the antigen target, rinsing and washing in buffers to minimize non-specific reactions, application of a secondary antibody specific to the primary antibody species conjugated to a tag (such as biotin or horseradish peroxidase), and finally addition of a detection reagent (such as a chromagen or fluorophore) to visualize the deposition of the primary antibody. A more detailed description of this technique can be found in a recent review by Taylor et al.[14] Because these studies can be performed on tissue that has been processed, fixed, and paraffin embedded in a standard fashion,

it can be utilized on biopsies or surgical specimens not only prospectively collected on clinical trials, but also on specimens obtained pre or post a specific treatment, and on archived samples.

Although both frozen and formalin-fixed paraffin-embedded (FFPE) processed tissue are used, frozen tissue allows for native structures within the cell to be maintained for DNA, RNA, and protein. This makes it very appealing for those working with gene expression arrays. However, when doing IHC or ISH, the tissue can be harder to work with due to temperature maintenance requirements, therefore FFPE tissue is the most common source of tissue for these studies.[15] The relative pros and cons of different tissue sources are outlined in Table 2.

Although IHC is a widely used technique, several disadvantages exist specific to the use of FFPE tissue. One problem relates to fixation times and requirements for antigen retrieval. Formalin fixation occurs through a combination of penetration and cross-linking. The rate of penetration of formalin through tissue has thus commonly been reported to be approximately 1mm/hour and therefore varies from sample to sample depending on the size of tissue and the more speculative variable of cross-linking time.[16] Because molecular changes may occur (and in a patchy fashion) during this period, the reliability and reproducibility of certain immunohistochemically detected biomarkers may be compromised.[17] Nevertheless, there are several studies that have positively evaluated IHC results over different fixation times and TMA studies that have shown reproducibility in staining and outcome correlations when looking at tissue across multiple decades.[18, 19]

Another important variable related to IHC-based studies is specificity of antibody. The methods

for antibody validation are certainly not standardized for even common commercially available

antibodies but often involve utilizing cell lines with known or tested concentrations of the

antigen compared by immunoblot analysis and/or ELISA. For our AQUA/TMA based studies

(see below), we use these same cell lines to make 'cell blocks' to place on a TMA to serve as an

internal control as well as a normalization standard for slide-to-slide comparisons.[20] Although

studies with blocking peptides is helpful, it does not give us an antibody's true specificity or

information regarding signal to noise ratios. Of course, simply observing staining patterns and

localization, confirming reproducibility, and if available, correlating with expected clinical

parameters may also assist in assessing the specificity of an antibody.

Tissue Microarrays

As noted above, tissue microarrays complement the large scale genomic/proteomic discovery

approach of expression arrays by allowing the simultaneous analysis of DNA, RNA, or protein in

large numbers of samples per single experiment (as opposed to DNA arrays which look at large

numbers of gene products simultaneously in a test sample). By linking these data to relevant

outcome information, e.g. survival, these analyses can give insight into the clinical significance

of a given biomarker.

Although the concept of standardizing and streamlining immunohistochemistry techniques have

been previously reported,[21] Kononen et al.[22] first described a device for the construction of

TMAs that could be feasibly accessible to many labs. The bulk of the time spent in TMA construction is the collection of the appropriate paraffin-embedded 'donor' tissue blocks and identification of the area of tissue of interest (e.g. invasive tumor). The 'recipient' or 'master' arrays are then assembled by taking a core tissue specimen from hundreds of separate donor blocks (e.g. different patient tumor blocks) and re-embedding them into the recipient block. Typically, cores are 0.6 mm in diameter spaced at 0.7-0.8 mm, which allows up to a 1000 samples to be placed on a recipient block. Larger diameter cores can also be taken in certain instances (for example when tissue heterogeneity is expected to be greater), although this reduces the number of cores that can be taken from the donor block and that can be placed into the recipient block. Depending on the thickness of the samples, 100-200 5µm sections can be cut from the recipient block for transfer onto a standard glass slide using an adhesive tape transfer method. The resultant slide can then be analyzed for a variety of molecular targets at the DNA, RNA, or protein level. Redundant arrays can also be constructed by obtaining multiple cores from the donor blocks and placing them at identical coordinates in recipient blocks.

Because both cut slides as well as blocks may be subject to antigen oxidation and degeneration, some facilities store recipient blocks in sealed nitrogen chambers and coat cut slides in paraffin to minimize these effects.[23] In addition, because tissue blocks are three-dimensional structures that can change as more sections are cut, most facilities employ a quality control monitoring system (e.g. every 10th section stained with H&E to assess tissue representativity).

There are several advantages to TMAs. First and most significant is the amplification of tissue resources. A conventional block would be exhausted by 50-100 cuts and analysis of 50

antibodies on 250 specimens would require 12,500 slides. This approach to tissue analysis is a prohibitive task that also very quickly exhausts tissue resources. As an example using TMAs, up to 400 master blocks can be made from a 1 cm tissue section. These can each be cut as many as 200 times, allowing the evaluation of 80,000 unique reagents. Second, this allows the efficient organization and storage of archived tissue blocks in many pathology departments. At the Yale TMA Facility, radiofrequency identification tags are used in the tissue blocks for this purpose. Third, large numbers of different types of tissues (benign and malignant), xenograft tissues, cell lines, or recombinant proteins can be readily integrated into the arrays to serve as intra- and inter-slide reference controls. Only a limited amount of antibody and other reagents, similar to what is used for a whole section, are required. In addition, because hundreds of samples can be studied in one experiment, common variables that can affect reproducibility, such as antigen retrieval, reagent concentrations, and washing times, can be standardized.

A common criticism of TMAs relates to the issue of tissue and tumor heterogeneity; whether a small core is representative of the entire tumor. Indeed, this argument can be expanded to whole tissue sections and blocks of tumor, as large surgical resections are only semi-randomly sampled by the pathologist. Nevertheless, many investigators have shown concordance rates of approximately 95% between 2-4 0.6 mm TMA spots and whole sections for common biomarkers such as estrogen receptor (ER) and progesterone receptor (PR) in breast cancer.[18, 24] Furthermore, they were able to reproduce known clinico-pathologic correlations with the TMA-based studies.[24] Similar validation studies have been performed in numerous other tumors, including those felt to be more inherently heterogenous such as Hodgkin's lymphoma,[25] pancreatic carcinoma,[26] and soft tissue sarcomas,[27] and colorectal cancer.[28] Although TMAs are

best used as epidemiology based research tools to examine relative expression of molecular markers in large cohorts of patients, these studies suggest that diagnostic application to individual clinical patients may also be appropriate if used judiciously.

Analysis and Interpretation of IHC studies

Manual scoring

For both routine histologic whole sections and for TMAs, the most common method of "quantifying" protein expression on immunostained tissue is visualization of chromagen (brown stain) intensity by light microscopy by a pathologist. This is a popular method used by nearly all clinical diagnostic studies (e.g. ER in breast cancer, epidermal growth factor receptor) for a number of reasons. Most importantly, it is easily performed and analyzed by a standard immunolab and by a pathologist's trained eye. Moreover, it is compatible with hematoxylin as a counter stain to delineate morphologic context. Although the readout is often a binary variable (positive or negative), efforts at semi-quantitative scoring (H scores and Allred scores for ER and 0, 1, 2, 3 scores for Her2/neu immunostains in breast cancer) have been partially successful.[8, 29] This method has also been extensively and most commonly used to analyze TMA-based IHC studies.[30]

Automated analysis of chromagen-based IHC

Despite the many pros of pathologist-based scoring of immunostains, there are some drawbacks as well. Numerous studies have now demonstrated the fairly high intra- and inter-observer variability seen with these analyses, presumably due to the subjective nature of these readouts.[31]

Thus in recent years, numerous technologies have been described with the goal towards a more automated image acquisition of microscopy fields and TMA spots and the production of a continuous data set of IHC scores. Most of these systems are based again on acquisition of traditional chromagen (brown stain) images with modified bright field microscopy. Table 3 discusses some of the currently available systems for automated analysis, many of which are programmed with the ability to determine the location of tissue spots in tissue microarrays.

Automated analysis of fluorescence based IHC

Because protein concentrations can span many logs of expression, the utilization of a detection system that spans as great as possible dynamic range to mimic the biologic levels of the protein is important. Brown stains typically have low absorbance and thus have limited dynamic range. Fluorescence-based detection systems may have an advantage based on its wider dynamic range and multiplexing capabilities.[32] Although tissues do have endogenous autofluorescence, this confounding variable can be minimized with adequate controls, and careful attention to fluorophore chosen and antibody titer.

Our group at Yale has developed a technology called AQUA that uses an immunofluoresence-based detection system which allows increased sensitivity and dynamic range.[33] The unique aspect of this system is that rather than using morphology to define compartments, it uses molecular tags (e.g. cytokeratin for epithelium and DAPI for nucleus). Using a series of algorithms to subtract out-of-focus from in-focus images (RESA) as well as the molecular co-localization algorithms (PLACE), protein expression is then assessed on a continuous scale

within the compartment of interest (e.g. tumor cells) normalized to area (Figure 1). This technology has been applied to the study of a variety of biomarkers in numerous different cancers.[34-36] A recent study showed that AQUA-based quantitation was proportional to ELISA-based concentrations of HER2 in cell lines suggesting that *in situ* quantitation may be possible with maintenance of spatial information.[37] In addition, because a molecular tag is simply defined by a molecule with specificity for a defined/localized antigen, one can use the technology to study protein expression in other compartments (e.g. golgi, mitochondria, microvessels) as well as virtual compartments (expression of protein A within area defined by protein B).[38] Although not yet clinically realized, several studies have suggested that compartmental localization of certain markers may indeed be important in the study of biological function as well as in the classification of certain diseases.[39]

Despite these benefits, this technology is still dependent on important human elements, for example to localize the relevant areas of study (e.g. tumor) versus other benign areas. Many of the systems utilizing fluorescence may also be limited to single exposure times while imaging a single slide. Some programs are being upgraded and tested for utilizing individual exposure times per image and correcting for the difference in algorithms for determining quantitation afterwards. And of course the limitations discussed above for FFPE regarding fixation of tissue and antibody validation are also relevant for AQUA.

CONCLUSION

Measurement of biomarkers and their profiles may be a more precise surrogate marker for older more traditional measurements of tumor behavior such as stage and grade. Thus in some ways we have come full circle. The advent of molecular biologic techniques have led us to a better understanding of tumor behavior and the more recent progress in high throughput expression profiling studies have furthered our understanding of the alterations that tumors have on a nucleic acid level. The results of these studies can be furthered triaged into the most relevant biomarkers by validating the most promising biomarkers on separate validation sets with different methods. Ideally, they should ultimately be tested on patient samples in well-designed prospective clinical trials to prove or disprove their clinical utility.

FIGURE LEGENDS

Figure 1. Schematic of AQUA on a TMA spot

A fluorescent labeled anti-cytokeratin (CK) antibody functions as an epithelial tag and distinguishes tumor cells from stromal elements. The assumption is that these are tumor cells not benign because TMA spots are preselected for tumor rich areas. DAPI is used as a nuclear tag. Using an automated subtraction algorithm (RESA), out-of-focus information is subtracted to produce the non-nuclear (cytokeratin stain) and nuclear (DAPI) compartments. The CK image is converted into a binary tumor mask. Each pixel in the tumor mask is defined as either nuclear or non-nuclear (i.e. membrane/cytoplasmic) to define the respective sub-cellular compartments. Target is visualized ustilizing a Cy5 labeled secondary antibody. The target image also undergoes RESA and the intensity is quantified (total signal intensity / area of the compartment). A constructed image of the target localization shows that the target is predominantly red (membrane/cytoplasm) rather than blue (nuclear).

REFERENCES

1.   Golub TR, Slonim DK, Tamayo P , et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;286:531-537.

2.   Liotta L, Petricoin E. Molecular profiling of human cancer. *Nat Rev Genet* 2000;1:48-56.

3.   Sorlie T, Perou CM, Tibshirani R , et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869-10874.

4.   Sotiriou C, Piccart MJ. Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? *Nat Rev Cancer* 2007;7:545-553.

5.   Chuaqui RF, Bonner RF, Best CJ , et al. Post-analysis follow-up and validation of microarray experiments. *Nat Genet* 2002;32 Suppl:509-514.

6.   Taniguchi M, Miura K, Iwao H , et al. Quantitative assessment of DNA microarrays--comparison with Northern blot analyses. *Genomics* 2001;71:34-39.

7.   Rajeevan MS, Vernon SD, Taysavang N , et al. Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. *J Mol Diagn* 2001;3:26-31.

8.   Carlson RW, Moench SJ, Hammond ME , et al. HER2 testing in breast cancer: NCCN Task Force report and recommendations. *J Natl Compr Canc Netw* 2006;4 Suppl 3:S1-22; quiz S23-24.

9.   Hayes DF, Thor AD, Dressler LG , et al. HER2 and response to paclitaxel in node-positive breast cancer. *N Engl J Med* 2007;357:1496-1506.

10.   Espina V, Mehta AI, Winters ME , et al. Protein microarrays: molecular profiling technologies for clinical specimens. *Proteomics* 2003;3:2091-2100.

11.   Paweletz CP, Charboneau L, Bichsel VE , et al. Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene* 2001;20:1981-1989.

12.   van de Rijn M, Gilks CB. Applications of microarrays to histopathology. *Histopathology* 2004;44:97-108.

13.   Giltnane JM, Rimm DL. Technology insight: Identification of biomarkers with tissue microarray technology. *Nat Clin Pract Oncol* 2004;1:104-111.

14.   Taylor CR, Levenson RM. Quantification of immunohistochemistry--issues concerning methods, utility and semiquantitative assessment II. *Histopathology* 2006;49:411-424.

15.   Watanabe A, Cornelison R, Hostetter G. Tissue microarrays: applications in genomic research. *Expert Rev Mol Diagn* 2005;5:171-181.

16.   Werner M, Chott A, Fabiano A , et al. Effect of formalin tissue fixation and processing on immunohistochemistry. *Am J Surg Pathol* 2000;24:1016-1019.

17.   Mandell JW. Phosphorylation state-specific antibodies: applications in investigative and diagnostic pathology. *Am J Pathol* 2003;163:1687-1698.

18.   Camp RL, Charette LA, Rimm DL. Validation of tissue microarray technology in breast carcinoma. *Lab Invest* 2000;80:1943-1949.

19.   Shi SR, Liu C, Taylor CR. Standardization of immunohistochemistry for formalin-fixed, paraffin-embedded tissue sections based on the antigen-retrieval technique: from experiments to hypothesis. *J Histochem Cytochem* 2007;55:105-109.

20.   Pozner-Moulis S, Cregger M, Camp RL , et al. Antibody validation by quantitative analysis of protein expression using expression of Met in breast cancer as a model. *Lab Invest* 2007;87:251-260.

21.   Battifora H. The multitumor (sausage) tissue block: novel method for immunohistochemical antibody testing. *Lab Invest* 1986;55:244-248.

22.   Kononen J, Bubendorf L, Kallioniemi A , et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* 1998;4:844-847.

23.   DiVito KA, Charette LA, Rimm DL , et al. Long-term preservation of antigenicity on tissue microarrays. *Lab Invest* 2004;84:1071-1078.

24.   Torhorst J, Bucher C, Kononen J , et al. Tissue microarrays for rapid linking of molecular changes to clinical endpoints. *Am J Pathol* 2001;159:2249-2256.

25.   Garcia JF, Camacho FI, Morente M , et al. Hodgkin and Reed-Sternberg cells harbor alterations in the major tumor suppressor pathways and cell-cycle checkpoints: analyses using tissue microarrays. *Blood* 2003;101:681-689.

26.   Maitra A, Adsay NV, Argani P , et al. Multicomponent analysis of the pancreatic adenocarcinoma progression model using a pancreatic intraepithelial neoplasia tissue microarray. *Mod Pathol* 2003;16:902-912.

27.   Engellau J, Akerman M, Anderson H , et al. Tissue microarray technique in soft tissue sarcoma: immunohistochemical Ki-67 expression in malignant fibrous histiocytoma. *Appl Immunohistochem Mol Morphol* 2001;9:358-363.

28.   Jourdan F, Sebbagh N, Comperat E , et al. Tissue microarray technology: validation in colorectal carcinoma and analysis of p53, hMLH1, and hMSH2 immunohistochemical expression. *Virchows Arch* 2003;443:115-121.

29.   Ellis MJ, Coop A, Singh B , et al. Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for ErbB-1- and/or ErbB-2-positive, estrogen receptor-positive
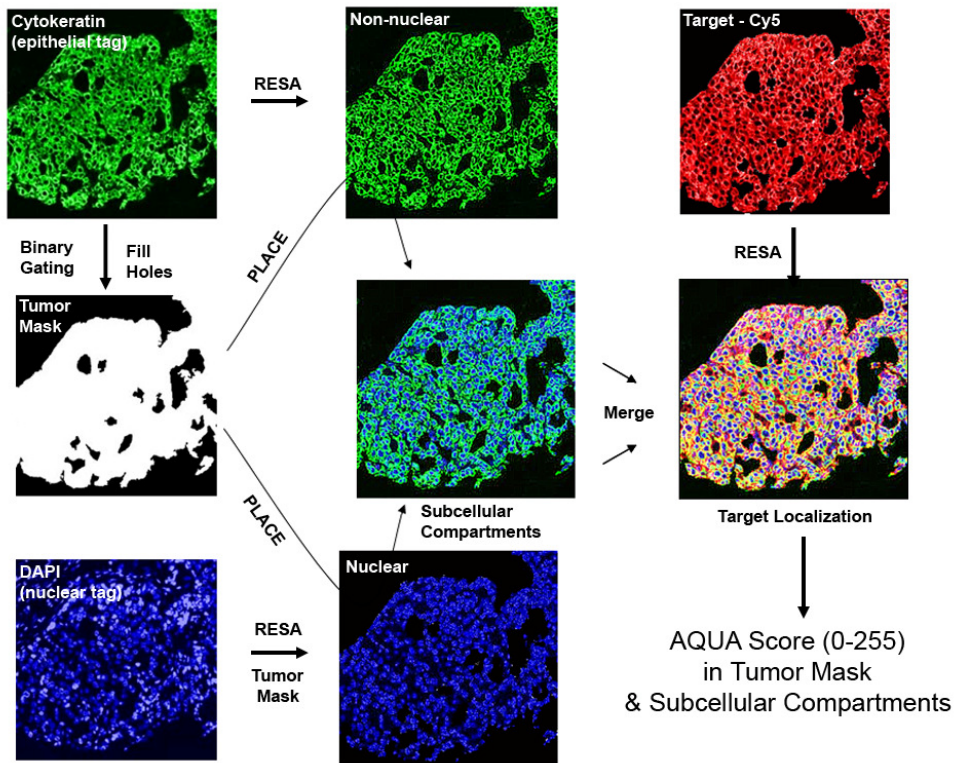
primary breast cancer: evidence from a phase III randomized trial. *J Clin Oncol* 2001;19:3808-3816.

30.   Chung GG, Kielhorn EP, Rimm DL. Subjective differences in outcome are seen as a function of the immunohistochemical method used on a colorectal cancer tissue microarray. *Clin Colorectal Cancer* 2002;1:237-242.

31.   Sirota RL. Error and error reduction in pathology. *Arch Pathol Lab Med* 2005;129:1228-1233.

32.   Rubin MA, Zerkowski MP, Camp RL , et al. Quantitative determination of expression of the prostate cancer protein alpha-methylacyl-CoA racemase using automated quantitative analysis (AQUA): a novel paradigm for automated and continuous biomarker measurements. *Am J Pathol* 2004;164:831-840.

33.   Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nat Med* 2002;8:1323-1327.

34.   Camp RL, Dolled-Filhart M, King BL , et al. Quantitative analysis of breast cancer tissue microarrays shows that both high and normal levels of HER2 expression are associated with poor outcome. *Cancer Res* 2003;63:1445-1448.

35.   Psyrri A, Yu Z, Weinberger PM , et al. Quantitative determination of nuclear and cytoplasmic epidermal growth factor receptor expression in oropharyngeal squamous cell cancer by using automated quantitative analysis. *Clin Cancer Res* 2005;11:5856-5862.

36.   Zerkowski MP, Camp RL, Burtness BA , et al. Quantitative analysis of breast cancer tissue microarrays shows high cox-2 expression is associated with poor outcome. *Cancer Invest* 2007;25:19-26.

37.   McCabe A, Dolled-Filhart M, Camp RL , et al. Automated quantitative analysis

(AQUA) of in situ protein expression, antibody concentration, and prognosis. *J Natl*

Cancer Inst 2005;97:1808-1815.

38.   Mertz KD, Demichelis F, Kim R , et al. Automated immunofluorescence analysis

defines microvessel area as a prognostic parameter in clear cell renal cell cancer. *Hum*

Pathol 2007;38:1454-1462.

39.   Berger AJ, Kluger HM, Li N , et al. Subcellular localization of activating transcription

factor 2 in melanoma specimens predicts patient survival. *Cancer Res* 2003;63:8103-

8107.

Table 1. *In situ* quantitative methods			
Method	**Description**	**Advantages**	**Disadvantages**
ISH	Use of labeled complimentary DNA or RNA strand to localize a specific DNA or RNA sequence of interest in tissue	Can help identify and determine structure of chromosomes, chromosomal integrity, localize and count amplified genes and mRNA transcripts	No information on protein levels, unclear sensitivity
IHC – Manual	Use of labeled antibodies to localize specific proteins of interest in tissue using chromagen based detection and bright field microscopy	Readily feasible for most clinical diagnostics labs	Limited to nominal (or semi-quantitative) scoring, influential to bias, small differences in intensity difficult to detect by human eye, antigen retrieval generally required, antibody specificity needs to be validated
IHC – Chromagen-based automation	Same as IHC-manual except automated measurement of protein based on intensity of staining	Continuous scoring	Limited to user-defined compartments, limited ability to multiplex several markers, lower dynamic range compared with fluorescence, antigen retrieval generally required, antibody specificity needs to be validated
IHC – Fluorescence-based automation	Use of labeled antibodies to localize specific proteins of interest in tissue, and automated measurement of protein based on emittance of fluorescent signal of target	Continuous scoring, ability to multiplex several markers	Autofluorescence, fewer validated studies, currently limited to research labs with limited commercial availability, antigen retrieval generally required, antibody specificity needs to be validated

Table 2. Sources of tissue for biomarker studies			
Source	**Description**	**Advantages**	**Disadvantages**
Frozen	Tissue sources fixed by acetone or OCT, then frozen by dry ice or flash frozen in liquid nitrogen	Higher sensitivity for ISH, more native structures due to less fixation of tissue that causes cross-linking	More difficult TMA construction – need to maintain temperature, morphology diminished
FFPE	Formalin-fixed paraffin-embedded tissue	Better morphology, works well with IHC and ISH	Must apply antigen retrieval, slow fixation time may cause changes of certain protein biomarkers and RNA levels
Whole Tissue Section	Section of tissue block removed and transferred onto slide for analysis	Standard diagnostics pathology technique, larger area than TMA for more variable and heterogeneous marker analysis	Can analyze only limited numbers of tissue sections at a time, little immediate normalization between tissue sections
TMA	Cores of multiple tissue blocks fixed in grids, then sectioned and fixed onto slide for analysis	Analysis of hundreds of samples simultaneously, ability to do large-scale analysis on biomarker, better standardization, amplification of tissue resources	Heterogeneity issues, not yet broadly available, not directly applicable to clinical diagnostic purposes

Table 3. Available automated analysis systems		
System (Manufacturer)	**Description**	**Commercially Available**
BLISS with TMAscore (Bacus Laboratories, Inc.)	Chromagen-based system, digital image acquisition of sections and TMAs, image analysis and quantitation using an intensity based algorithms	Yes
ACIS (Clarient)	Chromagen-based system, digital image acquisition, image analysis and quantitation using an intensity based algorithms (validated measurements of Her2/neu in breast cancer)	Yes
Pathiam Ruo (BioImagene)	Chromagen-based system, digital image acquisition, image analysis and quantitation, IHC algorithms for common biomarkers, morphological and intensity based algorithms, measures lengths and regions of interest	Yes
ScanScope (Aperio Technologies)	Chromagen-based system, digital image acquisition and quantitation of percent-positive cells, quantification of area and intensity of stain, compartmentalization algorithms	Yes
VIAS (Ventana Medical Systems, Inc.)	Chromagen-based system, real-time image capture and quantitation of percent-positive cells, marker intensity, size and shape	Yes
CMYK color model (Ontario Cancer Institute)	Chromagen-based system, digital image acquisition, image analysis and quantitation using an intensity based algorithms with yellow channel	No
iVision and GenMx Vision (Biogene	Chromagen- and Fluorescence-based system, digital image acquisition and quantitation of percent-positive cells, quantification of area and intensity of stain, compartmentalization algorithms	Yes
iColor and iCys Imaging Cytometer (CompuCyte)	Chromagen- and Fluorescent-based system, digital image acquisition, quantitation of marker intensity, co-localization to molecular-defined compartments, morphological imaging	Yes
Ariol SL-50 (Applied Imaging Corp.)	Chromagen- and Fluorescent-based system, digital image acquisition, quantification using intensity of stain and fluorescence in IHC and FISH, co-localization to compartments, multiple chromagenic channels, morphometric analysis	Yes
AQUA (HistoRx, Inc.)	Fluorescence-based system, utilizes a image grabber system to automate area of tissue section or spot location of TMA to image, determination of molecular-defined compartments, co-localization of target to compartments, ability of multiple targets	Yes

Cytokeratin (epithelial tag)

Binary Gating

Fill Holes

RESA

Non-nuclear

PLACE

Tumor Mask

Target - Cy5

RESA

PLACE

Subcellular Compartments

Merge

Target Localization

DAPI (nuclear tag)

RESA Tumor Mask

Nuclear

AQUA Score (0-255)
in Tumor Mask
& Subcellular Compartments

Quantitative analysis of estrogen receptor heterogeneity in breast cancer

Gina G Chung[1], Maciej P Zerkowski[2], Sriparna Ghosh[1], Robert L Camp[3] and David L Rimm[3]

Immunohistochemical analyses (IHC) of biomarkers are extensively used for tumor characterization and as prognostic and predictive measures. The current standard of single slide analysis assumes that one 5 μM section is representative of the entire tumor. We used our automated image analysis technology (AQUA) using a modified IHC technique with fluorophores to compare estrogen receptor (ER) expression in multiple blocks/slides from cases of primary breast cancer with the objective of quantifying tumor heterogeneity within sections and between blocks. To normalize our ER scores and allow slide-to-slide comparisons, 0.6 μm histospots of representative breast cancer cases with known ER scores were assembled into a 'gold standard array' (GSA) and placed adjacently to each whole section. Overall, there was excellent correlation between AQUA scores and the pathologist's scores and reproducibility of GSA scores (mean linear regression R value 0.8903). Twenty-nine slides from 11 surgical cases were then analyzed totaling over 2000 AQUA images. Using standard binary assignments of AQUA (>10) and pathologist's ($>10\%$) scores as being positive, there was fair concordancy between AQUA and pathologist scores (73%) and between slides from different blocks from the same cases (75%). However using continuous AQUA scores, agreement between AQUA and pathologist was far lower and between slides from different blocks from the same cases only 19%. Within individual slides there was also significant heterogeneity in a scattered pattern, most notably for slides with the highest AQUA scores. In sum, using a quantitative measure of ER expression, significant block-to-block heterogeneity was found in 81% of cases. These results most likely reflect both laboratory-based variability due to lack of standardization of immunohistochemistry and true biological heterogeneity. It is also likely to be dependent on the biomarker analyzed and suggests further studies should be carried out to determine how these findings may affect clinical decision-making processes.
Laboratory Investigation advance online publication, 5 March 2007; doi:10.1038/labinvest.3700543

It has long been recognized that breast cancer exhibits widely disparate clinical behavior that cannot be solely attributable to stage. Recent expression profiling studies have corroborated this finding by identifying biologically distinct groups of tumors spanning traditional classification schemas such as stage and tumor size.[1] Within an individual tumor, biologically relevant heterogeneity may also exist in part, owing to variations in the tumor microenvironment, cell cycle variations, and stem cell subpopulations.

Immunohistochemical (IHC) evaluation of biomarkers in tissue has evolved to become a commonly used diagnostic tool for the pathologist, predominantly because of its ease and ability to retain morphologic information. In addition, tissue microarrays (TMA) have become a popular tool for the rapid and efficient detection of clinicopathologic associations in large numbers of samples.[2,3] Major criticisms of TMAs relate to tumor heterogeneity and whether small TMA cores are representative of the whole section. On a broader scale, these criticisms are limited because they are based on the assumption that whole sections accurately reflect the entire tumor. For example, tissue is typically sampled at the rate of one section per cm^3 of tumor. Although the volume of each TMA spot represents only about 0.0002% ($0.6 \times 0.6 \, mm^2$ diameter and 5 μm thick) of this tumor, the standard tissue section would represent only 0.05% ($1 \times 1 \, cm$ dimension and 5 μm thick) of this tumor.

Estrogen receptor (ER) is an important regulator of both physiologic and pathologic mammary growth and differentiation.[4,5] Although its expression has been associated with well differentiated, lower grade tumors, most studies have

[1]Medical Oncology, Yale Cancer Center, New Haven, CT, USA; [2]HistoRx Inc., New Haven, CT, USA and [3]Department of Pathology, Yale University School of Medicine, New Haven, CT, USA
Correspondence: Dr GG Chung, MD, Medical Oncology, Yale Cancer Center, 333 Cedar Street, PO Box 208032, New Haven, CT 06520, USA. E-mail: gina.chung@yale.edu

shown that it also retains independent prognostic information.[6] More importantly, ER expression is the most reliable predictor of response to endocrine therapies in breast cancer.[7] Current standard of care for determining ER status of a given clinical sample is IHC on a whole section with manual readouts. Although individual tumors typically have relatively homogenous ER expression, staining heterogeneity has been observed in some tumors.[8]

We have developed an algorithm for quantitatively determining *in situ* protein expression called automated image analysis technology (AQUA).[9] AQUA is a hybrid of standard IHC and flow cytometry in that it requires antigen retrieval on fixed tissue, application of primary and secondary antibodies, and use of multiplexed fluorescent detection to produce an objective, numeric score. This technology reduces the bias of subjective assessment and allows quantification of protein expression using molecular colocalization techniques. Unlike either technology however, the resultant score is directly proportional to the concentration of the target protein within a user-defined compartment.[10] This methodology has been validated with a variety of biomarkers in many different cancers.[11–13] Thus, we used AQUA to collect continuous ER scores from multiple images/fields of standard whole-tissue sections from multiple blocks of the same surgical cases with the objective of quantifying tumor heterogeneity.

MATERIALS AND METHODS
Case Selection
Eleven cases of primary invasive ductal adenocarcinomas of the breast from 1999 to 2000 were retrieved from the Yale University Pathology archives. Specifically, cases were selected that spanned a broad range of ER expression as determined by the pathologist at routine signout (eg 0–100%), that had abundance of tumor present in the blocks, and that had multiple blocks available from the same surgical case. Areas of invasive tumor were identified by a pathologist and circled on the whole section, giving careful attention to avoid areas with admixed *in situ* and/or benign tissue. Overall, there were 29 slides from 11 cases (either one, two, or three blocks from the same case and one section studied per each block). This study was approved by the Yale University Human Investigations Committee.

Gold Standard Array/Control Array
In order to normalize our ER scores and allow slide-to-slide comparisons, a 'gold standard array' (GSA) of exemplary, representative breast cancer cases with known ER scores was constructed. Six cases of breast cancer from 2002 were selected from the archives of the Yale University Department of Pathology. These cases also had abundant invasive cancer in the blocks and had a relatively homogenous appearance to the tumor. The signout pathologist's ER scores were 0, 20, 50, 70, 80, and 100% and the pathologists in this study felt these were classic or exemplary examples of these scores. Representative areas of invasive tumor were identified by

a pathologist and 0.6 mm diameter cores were placed into a recipient block using a precision arraying instrument (Beecher Instruments, Silver Spring, MD, USA). First, a master/triplicate GSA was assembled by taking three consecutive cuts (5 μm) of the GSA recipient block and affixing to an adhesive slide using a UV crosslinkable tape transfer system. Then, one 5 μm cut of the GSA was affixed onto slides adjacent to each whole section from the 29 slides described above to be stained and analyzed concurrently.

Immunohistochemistry
Staining slides for AQUA has been previously described.[9] Briefly, slides were deparaffinized in xylene, rinsed in ethanol, and rehydrated. Antigen retrieval was performed by pressure cooking for 15 min in 6.5 mM sodium citrate buffer. Endogenous peroxidase was quenched by immersing the array in a 2.5% methanol/hydrogen peroxide buffer for 30 min. Nonspecific background staining was further minimized by preincubating the array with 0.3% bovine serum albumin in 0.1 M Tris-buffered saline (pH 8.0) for 1 h. Primary antibodies used were pancytokeratin and ER, clone 1D5 (DAKO, Carpinteria, CA, USA). This primary antibody cocktail was incubated overnight at 4°C in a humidity chamber. Goat anti-mouse antibody conjugated to a horseradish peroxidase-decorated dextran polymer backbone (Envision; DAKO Corp.) was used as a secondary reagent to detect the bound primary target (ER) and Cy5-tyramide was used to visualize the amplified signal. Cy-5 (red) was used because its emission peak is well outside the green-orange spectrum of tissue autofluorescence. The cytokeratin was visualized with a Cy3-conjugated secondary antibody and the array was then counterstained with 4′,6-diamidino-2-phenylindole (DAPI) to define the nuclear compartment.

Image Collection and AQUA Analysis
Image acquisition and automated analysis have also been described extensively in previous work.[9] For the whole section analysis, images for the GSA and whole sections on each slide were captured separately. Images are automatically acquired with a high-resolution monochromatic camera using filter cubes specific to the emission/excitation spectra of DAPI, Cy5, and Cy3. Then, using this stack of uncompressed images, the AQUA software then allows one to distinguish between areas of tumor and stromal elements using the cytokeratin stain, resulting in a unique binary cytokeratin tumor mask for each spot. Furthermore, the cytokeratin and DAPI stains are used to assign each pixel under the tumor mask into non-overlapping membrane/cytoplasmic and nuclear locales. AQUA scores for ER are then calculated that correspond to the average signal intensity divided by compartment area. Although non-nuclear ER expression may be biologically relevant, as standard ER analysis relies on nuclear expression, we chose to quantitate ER signal in the nuclear compartment, that is, AQUA ER score within the DAPI-defined nuclear compartment. The AQUA score is thus pro-

portional to the average protein concentration in the nucleus across all of the cells within the keratin staining mask. This information can then be exported in a format suitable for analysis by standard statistical software packages.

For the whole sections, using a ×20 objective, a series of image 'fields' were captured within the circled invasive tumor to ultimately cover the vast majority of tissue of interest. Depending on the size of the tumor, 33–147 fields were captured per section and a total of 2043 fields were analyzed. Postimage capture, images were reviewed, particularly near the edges, to ensure that only fields within circled area of invasive tumor were included for analysis.

RESULTS
Analysis of GSA and Normalization

Three consecutive 5 μm cuts of the GSA were placed on the same slide, stained with cytokeratin, ER, and DAPI and analyzed with AQUA. Figure 1a shows the mean of the master/triplicate AQUA scores compared with the pathologist's ER scores. Overall, there was an excellent correlation between the mean AQUA scores and the pathologist scores. However, because the AQUA scores represent ER concentration, and whereas the pathologist scores represent simply the percent nuclei positive without regard to signal intensity or concentration, the relationship is not strictly linear.

Next, log AQUA scores of the GSA histospots on each whole-section slide was compared with the mean log scores on the master/triplicate GSA. Of note, identical staining conditions were used for all slides, including antibody concentrations, incubation times, etc. Figure 1b shows a representative case with linear regression $R = 0.8119$. For all 29 slides, linear regression R values ranged between 0.7750 and 0.9853, mean 0.8903. These graphs were used to normalize the individual raw AQUA scores from each field on the whole sections.

Inter-Slide Comparisons

Figure 2 shows slide-to-slide comparisons matched against the signout pathologist's score. Interslide differences appear to be minimized with the lowest scores. In our experience,

AQUA ER scores greater than ten are typically considered positive associated with high pathologic scores and with better prognosis in TMA studies.[14] Similarly, a pathologist's score of greater than 10% is traditionally considered positive. Using these *a priori* divisions to create binary values for AQUA and the pathologist's scores, the concordancy between AQUA and pathologist is reasonably good at 73% (κ score = 0.526). Furthermore, the three discordant cases (F, G, and J) all show individual fields above the threshold for positive cases. If these field were selected (instead of the average) the concordance would be 100%. Table 1 shows normalized, mean AQUA scores for the different blocks. Again, using the binary values for AQUA ER scores, the concordancy between different blocks from the same case is 75% (24 out of a total of 32 block-to-block comparisons concordant). Of the comparisons that were considered discordant, the absolute differences were small (eg Case J/Block
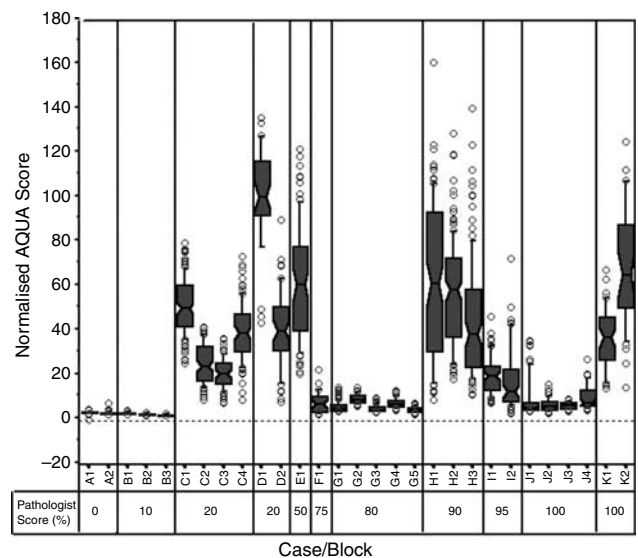

Figure 2 Box-plots of AQUA ER scores compared with pathologist's scores. Normalized AQUA ER scores are compared with the pathologist's scores for different blocks/slides for each case (eg A1 and A2 are two different blocks/whole sections from Case A).
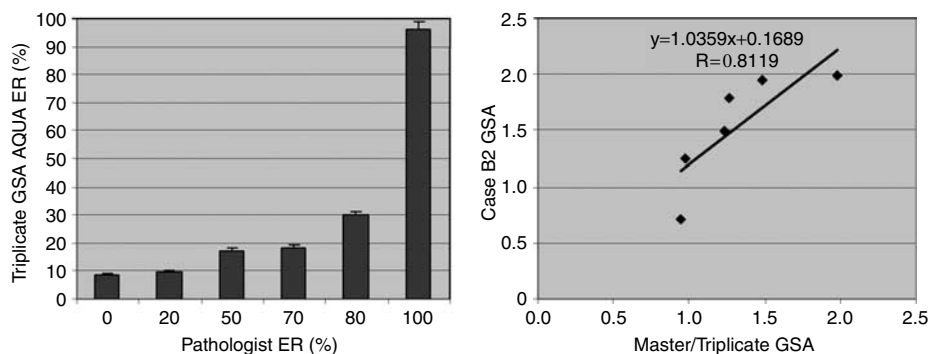

Figure 1 Gold standard array analysis. AQUA ER scores on three-fold redundant TMAs (triplicate GSA) were averaged and compared with the pathologist's ER estimation (**a**). These scores were then compared with the GSA scores (same cases as those on the triplicate GSA) on each whole-section analyzed to serve as internal controls and as a means of normalizing each whole section raw AQUA scores with the Master/Triplicate GSA.

Table 1 Slide-to-slide ER heterogeneity

Case	Block	Pathologist (%)	Mean AQUA[a]	Standard error	P value; Block comparisons	P value; ANOVA
A	1	0	3.667	0.074	0.2446; 1 *vs* 2	0.1995
	2		3.461	0.152		
B	1	10	3.332	0.072	<0.0001; 1 *vs* 2	<0.0001
	2		2.911	0.039	<0.0001; 2 *vs* 3	
	3		2.51	0.048	<0.0001; 1 *vs* 3	
C	1	20	51.803	1.105	<0.0001; 1 *vs* 2	<0.0001
	2		25.884	0.976	<0.0001; 1 *vs* 3	
	3		21.433	0.733	<0.0001; 1 *vs* 4	
	4		40.12	1.444	0.0077; 2 *vs* 3	
					<0.0001; 2 *vs* 4	
					<0.0001; 3 *vs* 4	
D	1	20	100	2.55	<0.0001; 1 *vs* 2	<0.0001
	2		40.978	3.71		
G	1	80	6.356	0.228	<0.0001; 1 *vs* 2[b]	<0.0001
	2		10.063	0.289	0.0018; 1 *vs* 3	
	3		5.527	0.138	<0.0001; 1 *vs* 4	
	4		8.105	0.225	<0.0001; 1 *vs* 5	
	5		4.979	0.167	<0.0001; 2 *vs* 3[b]	
					<0.0001; 2 *vs* 4[b]	
					<0.0001; 2 *vs* 5[b]	
					<0.0001; 3 *vs* 4	
					0.1023; 3 *vs* 5	
					<0.0001; 4 *vs* 5	
H	1	90	63.836	3.31	<0.0001; 1 *vs* 3	<0.0001
	2		57.639	1.935	0.0811; 1 *vs* 2	
	3		45.614	2.466	0.0007; 2 *vs* 3	
I	1	95	20.845	1.027	0.2047; 1 *vs* 2	0.2047
	2		18.306	1.756		
J	1	100	10.151	1.162	0.0042; 1 *vs* 2[b]	0.0253
	2		7.035	0.366	0.0117; 1 *vs* 3[b]	
	3		6.93	0.285	0.69; 1 *vs* 4	
	4		10.614	0.797	0.9348; 2 *vs* 3	
					0.0025; 2 *vs* 4[b]	
					0.0067; 3 *vs* 4	
K	1	100	36.778	2.056	<0.0001; 1 *vs* 2	<0.0001
	2		69.955	4.36		

[a]Normalized score to tissue controls and to maximum score (case D1).
[b]Discordant block to block comparisons with AQUA binary values.

1 *vs* Case J/Block 2 had AQUA scores of 10.151 and 7.035, respectively) (Table 1).

Most notable, however, are the discordancies between the pathologist and AQUA for cases F, G, J, and to a lesser degree case I (Figure 2 and Table 1). We attempted to retrieve the original ER slides assessed by the signout pathologist for these cases, but owing to the age of the research cases, we were only able to locate case I. Our review of this slide showed that the tumor appeared to be diffusely but weakly stained in a fairly nonspecific pattern (Figure 3a). There were however scattered areas of strong nuclear staining in adjacent benign components (Figure 3b). Because AQUA gives the average signal intensity in all pixels in a molecularly defined compartment (ie nucleus), it is possible that this may have accounted for the discordancy in this particular case. This case highlights the advantages of quantitative analysis when coupled to molecular compartment colocalization.

Because AQUA gives us continuous scores, we then performed unpaired *t*-tests and ANOVA analysis between the normalized, mean AQUA scores from different slides within individual cases (Table 1). Contrary to our results using binary divisions, this showed that only 6 of 32 (19%) slide-to-slide comparisons were concordant (*t*-test comparison *P*-value > 0.05). Only cases A and I, 2 of 9 (22%), were concordant (ANOVA *P*-value > 0.05). Case B was discordant with all three slide-to-slide comparisons significantly different. However, the AQUA scores were all extremely low (< 5) making these cases likely 'ER negative,' and the significance of this discordancy unclear. In addition, this decreased our concordancy with the pathologist's score. At first, this seems contrary to our findings on our master/triplicate GSA in which we saw good correlation between AQUA scores and pathologist's scores (Figure 1). However, the 0.6 mm diameter histospots on the GSA were carefully chosen as the most homogeneous appearing areas of the tumor with likely more homogeneous ER expression, and multiple blocks from these cases were not evaluated. For the 29 whole sections, many fields from each section and multiple blocks from each case were analyzed, thus maximizing apparent tumor heterogeneity. These data confirm the fact that field selection for scoring can dramatically change outcome.

2-D Spatial Heterogeneity

To further characterize heterogeneity within individual slides, we looked at the pattern of ER expression on each slide. Our normalized AQUA ER scores ranged from 2.959 to 174.672. Most of the slides with low AQUA scores (≤ 10) were relatively tightly clustered with minimal variance (Figure 4a). However, as the scores on a given slide increased, the variance generally increased (Figure 4a). This finding did not seem to be strictly related to the number of fields analyzed per slide as high variance was seen with high number of fields (Figure 4a, Case C) as well as with low number of fields (Figure 4a, Case K). Corresponding 2-D 'heat maps' were also generated based on the normalized AQUA scores (Figure 4b). The heat maps show the relative score of each field with respect to all of the other fields on the same slide. As is traditionally done for array illustration, increasing red intensity correlates with concentrations of ER above the mean and increasing green intensity correlates with degree below the mean. Although for most slides, high and low scores on a given section appeared to be randomly scattered and with a normal distribution
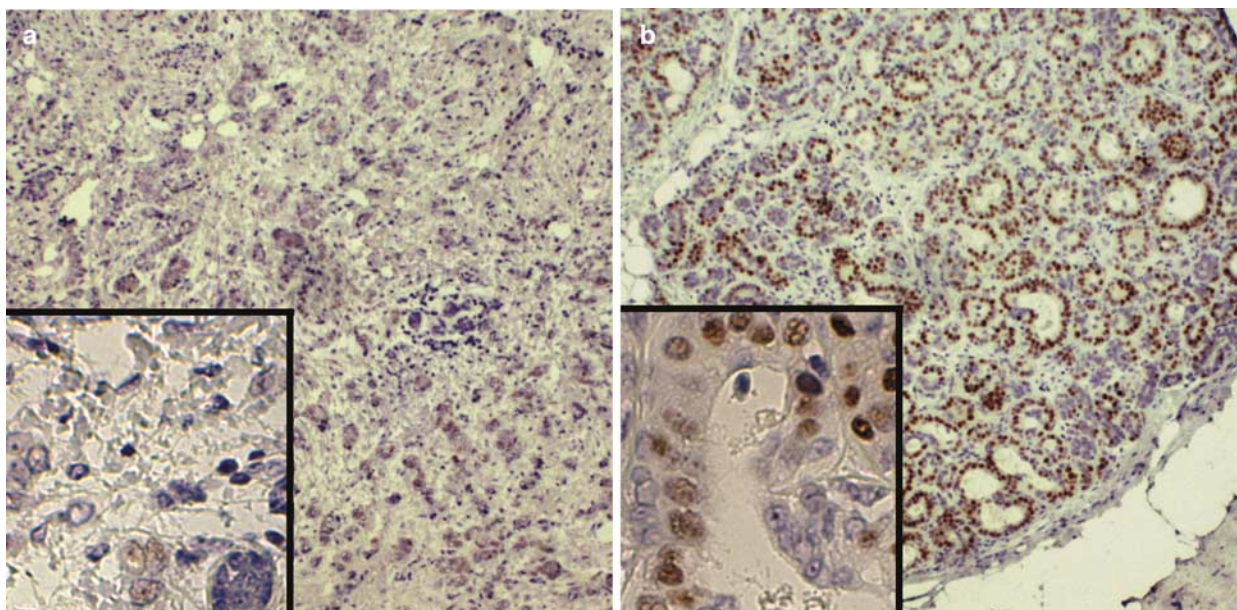

Figure 3 Case I used by the pathologist to generate the clinical ER score. (**a**) low-power image of the ER staining in the invasive tumor is shown. (**b**) One of several areas on the slide showing strong nuclear staining in an adjacent benign lesion is shown.
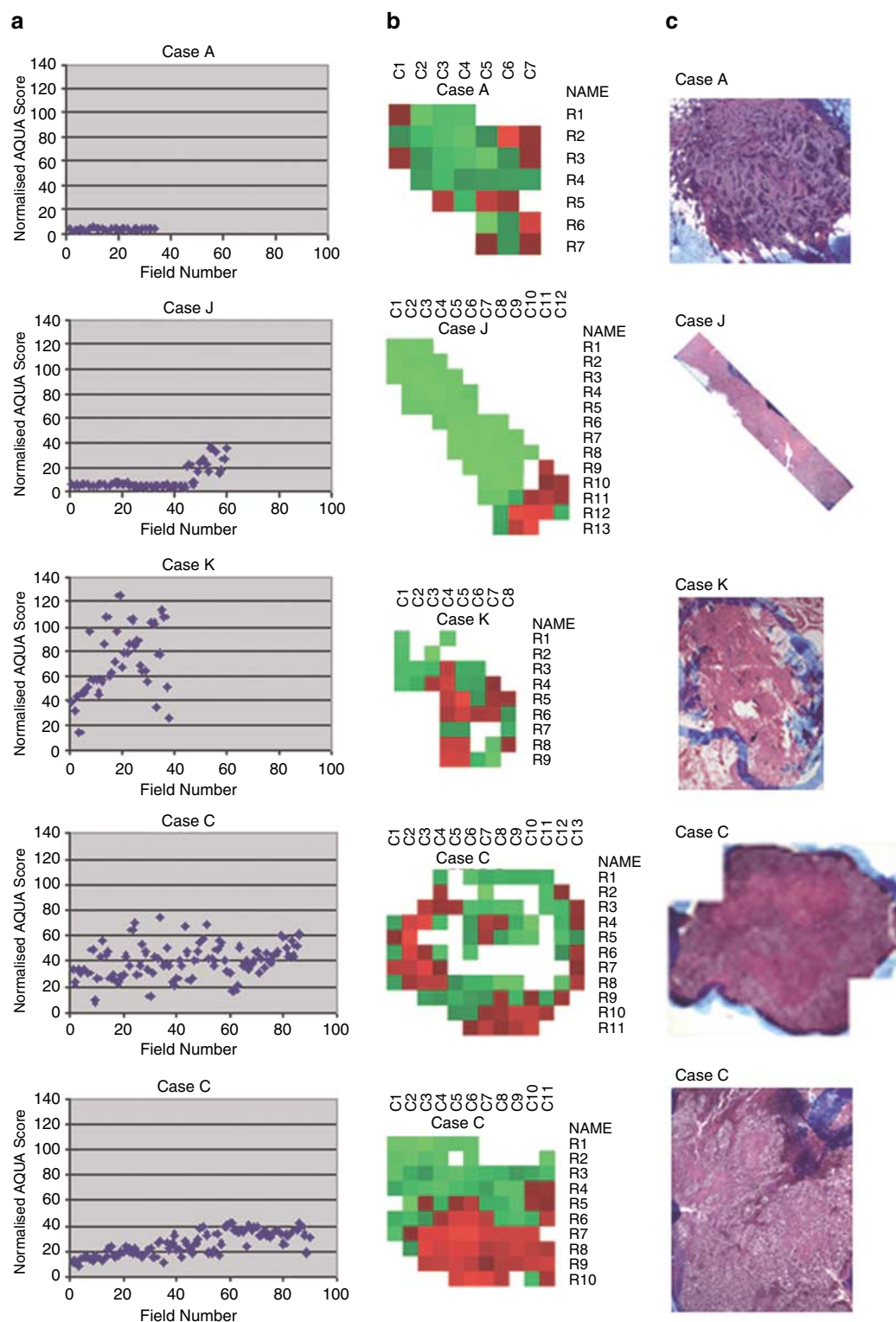
Figure 4 Scatter plots and 2-D representations of whole sections. (**a**–**c**) Representative data from five matched whole sections. (**a**) scatter plots of normalized AQUA scores of estrogen receptor are shown. Each image obtained on the section is assigned a field number starting in the upper left corner of the slide, across the row and down to the next row in a serpentine pattern. (**b**) heat maps showing AQUA ER scores assembled into a 'virtual slide' are shown. The most intense red are the highest ER scores and the most intense green are the lowest ER scores. (**c**) Corresponding H&E-stained whole sections used to identify areas of invasive cancers.

throughout the tumor, several slides showed a clustered pattern (Figure 4b, Cases J and 4b, and C). Interestingly, this clustering was seen in low scoring 'ER negative' cases as well as in higher scoring 'ER positive' cases. Indeed, scattered as well as clustered patterns were seen even on different blocks from the same case (Figure 4b, Case C). Corresponding whole sections stained with hematoxylin and eosin are also shown (Figure 4c) to demonstrate the areas of invasive cancers analyzed and the non-evaluable areas (shown as white squares in Figure 4b), which in general were areas without any tissue, minimum 'maskable' tumor, and areas having predominantly stroma or necrosis.

DISCUSSION

In this report, we have used a quantitative measure of ER expression, AQUA, to demonstrate significant block-to-block heterogeneity of IHC reaction. Our results most likely reflect both laboratory-based variability owing to lack of standardization of immunohistochemistry and true biological heterogeneity. These results raise several questions. Whereas the AQUA algorithm allows separation of epithelial cells from stroma with the keratin mask, contamination of the analysis with benign and/or in situ epithelial cells remain a possibility. These confounding factors were minimized because, as described above, analysis was limited as much as possible to previously identified areas of invasive cancer, and because adjacent in situ tumors typically show similar ER expression compared with the invasive component.[15] In addition, a number of investigators have described more intense staining in the periphery of tumors compared with the center.[16] This can be attributed in part to increased necrosis in the center (these areas are eliminated from AQUA analysis with a crop function) and fixation artifact. Some have reported this pattern more in surgical specimens than in matched core biopsies, suggesting that improper fixation of the central tumor in large specimens may play a role.[16] Several of our slides such as in (Figure 4b), Cases A and J also suggest more intense staining in the periphery. Alternatively, many investigators have described a phenomenon in which tumor at the invasive front shows different morphologies and preferentially expresses certain proteins vs tumor at the trailing edge/center.[17] Although this has been most extensively described for colorectal cancers, this has also been described in breast cancer in the so-called prairie fire pattern.[18] Irrespective of the reasons for the heterogenous pattern of ER expression within individual slides when using quantitative assessments, it is not surprising that there is a significant block-to-block heterogeneity within cases when using similar quantitative measurements.

At first, our concordance rate appears lower than previous studies looking at different assays for ER in the same tumor, ER assessments in matched core biopsies and surgical resections, and comparative ER levels in matched primary tumors and their metastases.[16,19–21] However, earlier studies that used biochemical assays such as the dextran-coated

charcoal method showed remarkably similar levels of ER discordancy within tumors ranging from 17 to 40% when using a binary cutpoint for ER positivity (eg ≥ 10 fmol/mg).[22–24] When the differences in the mean content were compared within tumors however, the discordancy rate was significantly higher. Our results with AQUA using binary divisions and then continuous measurements mimic closely the results seen in these prior studies. This suggests that our methodology may have the capacity to combine the practicality, morphologic information achievable with an in situ assay, and arguably superior predictive capacity of IHC assessment of ER measurements with the continuous readouts of ER levels analogous to a biochemical assay. This may also be important because higher levels of ER expression both by biochemical methods and by semiquantitative IHC readings (eg Allred score) have been associated with a greater likelihood of endocrine therapy response.[7,25] Ultimately, the value of an exact ER score may be seen in the context of other exact scores like PR) in distinguishing among various hormonal therapy options.

In the future, it is conceivable that other potential prognostic and predictive biomarkers may also rely more heavily on continuous readouts and more accurate assessments of total tumor heterogeneity. For example, if the division for optimally choosing patients for a given biologic therapy is between the highest expressers vs the high expressers (eg AQUA <150 vs >150), simply looking for evidence of minimal staining (none vs any) before choosing to give treatment (as is often done for ER and tamoxifen treatment) and more cautious assessment of tumor heterogeneity for that particular biomarker will be necessary.

In summary, using AQUA to quantify in situ ER expression on multiple blocks from different primary breast cancers, we demonstrate reasonable correlation using traditional binary divisions, although significant slide-to-slide tumor heterogeneity was seen in the majority of cases when continuous scores were analyzed. Heterogeneity was greatest for those with higher levels of expression. It is likely that tumor heterogeneity is highly dependent on biomarker analyzed and caution should be used when making IHC determinants of biomarker expression in a tumor with single slide assessments.

1. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA 2001;98:10869–10874.

2. Giltnane JM, Rimm DL. Technology insight: identification of biomarkers with tissue microarray technology. Nat Clin Pract Oncol 2004;1: 104–111.

3. Kononen J, Bubendorf L, Kallioniemi A, *et al*. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 1998;4:844–847.

4. Anderson E, Clarke RB, Howell A. Estrogen responsiveness and control of normal human breast proliferation. J Mammary Gland Biol Neoplasia 1998;3:23–35.

5. Sommer S, Fuqua SA. Estrogen receptor and breast cancer. Semin Cancer Biol 2001;11:339–352.

6. Fisher B, Redmond C, Fisher ER, *et al*. Relative worth of estrogen or progesterone receptor and pathologic characteristics of differentiation as indicators of prognosis in node negative breast cancer patients: findings from National Surgical Adjuvant Breast and Bowel Project Protocol B-06. J Clin Oncol 1988;6:1076–1087.

7. Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomised trials. Lancet 1998;351:1451–1467.

8. Layfield LJ, Saria E, Mooney EE, *et al*. Tissue heterogeneity of immunohistochemically detected estrogen receptor. Implications for image analysis quantification. Am J Clin Pathol 1998;110: 758–764.

9. Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. Nat Med 2002;8:1323–1327.

10. McCabe A, Dolled-Filhart M, Camp RL, *et al*. Automated quantitative analysis (AQUA) of *in situ* protein expression, antibody concentration, and prognosis. J Natl Cancer Inst 2005;97:1808–1815.

11. Camp RL, Dolled-Filhart M, King BL, *et al*. Quantitative analysis of breast cancer tissue microarrays shows that both high and normal levels of HER2 expression are associated with poor outcome. Cancer Res 2003;63:1445–1448.

12. Chung GG, Yoon HH, Zerkowski MP, *et al*. Vascular endothelial growth factor, FLT-1, and FLK-1 analysis in a pancreatic cancer tissue microarray. Cancer 2006;106:1677–1684.

13. Psyrri A, Yu Z, Weinberger PM, *et al*. Quantitative determination of nuclear and cytoplasmic epidermal growth factor receptor expression in oropharyngeal squamous cell cancer by using automated quantitative analysis. Clin Cancer Res 2005;11: 5856–5862.

14. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res 2004;10:7252–7259.

15. Leong AS, Sormunen RT, Vinyuvat S, *et al*. Biologic markers in ductal carcinoma *in situ* and concurrent infiltrating carcinoma. A comparison of eight contemporary grading systems. Am J Clin Pathol 2001;115:709–718.

16. Douglas-Jones AG, Collett N, Morgan JM, *et al*. Comparison of core oestrogen receptor (ER) assay with excised tumour: intratumoral distribution of ER in breast carcinoma. J Clin Pathol 2001;54:951–955.

17. Ono M, Sakamoto M, Ino Y, *et al*. Cancer cell morphology at the invasive front and expression of cell adhesion-related carbohydrate in the primary lesion of patients with colorectal carcinoma with liver metastasis. Cancer 1996;78:1179–1186.

18. Kauppila S, Stenback F, Risteli J, *et al*. Aberrant type I and type III collagen gene expression in human breast cancer *in vivo*. J Pathol 1998;186:262–268.

19. Allred DC, Bustamante MA, Daniel CO, *et al*. Immunocytochemical analysis of estrogen receptors in human breast carcinomas. Evaluation of 130 cases and review of the literature regarding concordance with biochemical assay and clinical relevance. Arch Surg 1990;125:107–113.

20. Mann GB, Fahey VD, Feleppa F, *et al*. Reliance on hormone receptor assays of surgical specimens may compromise outcome in patients with breast cancer. J Clin Oncol 2005;23:5148–5154.

21. Spataro V, Price K, Goldhirsch A, *et al*. Sequential estrogen receptor determinations from primary breast cancer and at relapse: prognostic and therapeutic relevance. The International Breast Cancer Study Group (formerly Ludwig Group). Ann Oncol 1992;3:733–740.

22. Davis BW, Zava DT, Locher GW, *et al*. Receptor heterogeneity of human breast cancer as measured by multiple intratumoral assays of estrogen and progesterone receptor. Eur J Cancer Clin Oncol 1984;20:375–382.

23. Lluch A, Anton A, Carbonell F, *et al*. Intratumoral variations in cytoplasmic estrogen receptor levels in breast cancer. Breast Cancer Res Treat 1987;9:69–70.

24. Silfversward C, Skoog L, Humla S, *et al*. Intratumoral variation of cytoplasmic and nuclear estrogen receptor concentrations in human mammary carcinoma. Eur J Cancer 1980;16:59–65.

25. Ellis MJ, Coop A, Singh B, *et al*. Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for ErbB-1- and/or ErbB-2-positive, estrogen receptor-positive primary breast cancer: evidence from a phase III randomized trial. J Clin Oncol 2001;19:3808–3816.